BMSAP

Bulletins et mémoires de la Société d'Anthropologie de Paris

33 (1) | 2021
Les 1000 premiers jours de vie dans les populations du présent et du passé

Analytical techniques and software for the study of intragroup metric variation using principal component analysis

Techniques analytiques et logiciels pour l'étude de la variation métrique intragroupe à l'aide de l'analyse en composantes principales

Andrej Alexeevich Evteev, Nikolai Evgenievich Staroverov and Nikolai Nikolaevich Potrakhov

OpenEdition Journals

Electronic version
URL: https://journals.openedition.org/bmsap/7539
DOI: 10.4000/bmsap.7539
ISSN: 1777-5469

Publisher
Société d'Anthropologie de Paris

Electronic reference
Andrej Alexeevich Evteev, Nikolai Evgenievich Staroverov and Nikolai Nikolaevich Potrakhov, "Analytical techniques and software for the study of intragroup metric variation using principal component analysis", *Bulletins et mémoires de la Société d'Anthropologie de Paris* [Online], 33 (1) | 2021, Online since 27 April 2021, connection on 04 June 2021. URL: http://journals.openedition.org/bmsap/7539 ; DOI: https://doi.org/10.4000/bmsap.7539

Analytical techniques and software for the study of intragroup metric variation using principal component analysis

Techniques analytiques et logiciels pour l'étude de la variation métrique intragroupe à l'aide de l'analyse en composantes principales

Andrej Alexeevich Evteev [1]*, Nikolai Evgenievich Staroverov [2], Nikolai Nikolaevich Potrakhov [2]

1 Anuchin Research Institute and Museum of Anthropology, Lomonosov Moscow State University, Moscow, Russia
2 Saint Petersburg Electrotechnical University "LETI", Department of Electronic Instruments and Devices, St.-Petersburg, Russia

* evteandr@gmail.com

Reçu : 4 août 2020 ; accepté : 5 février 2021
Bulletins et Mémoires de la Société d'Anthropologie de Paris

Abstract – Intragroup variation in human cranial samples is much less well understood than intergroup variation. The aims of this study were to develop a flexible and assumption-free approach for detailed explorations and comparisons of intragroup metric variation in any number of samples and to create user-friendly software for these purposes. We revisited the classic study design based on a comparison between the samples from Berg, Zalavar and Oslo from the W.W. Howells craniometric data set. Fourteen mid-facial dimensions were chosen for the analysis. "WorldPCA" software was employed for most of the analyses. This software implements a number of analytical functions aimed at exploring the results of principal component analysis. Our results confirm that the male sample from Berg displays a higher degree of variation. The cluster analyses have shown that intragroup variation in the three samples is mainly of a continuous nature. Arguably, the tendency to separate into distinct clusters is more pronounced in the samples from Oslo and Berg than from Zalavar. Some male individuals from Zalavar display distinct craniofacial features similar to those found in a South Siberian sample. Potential applications of these techniques and software are not restricted to cranial measurements but can be used for exploring any type of continuously varying data. No assumptions about the nature of the data should be made, and any number of samples can be compared simultaneously.

Keywords – intragroup variation, principal component analysis, Howells data set

Résumé – La variation intragroupe dans les échantillons crâniens humains est beaucoup moins bien comprise que la variation intergroupe. Cette étude avait pour objectif de développer une approche flexible et sans hypothèse pour l'exploration détaillée et la comparaison de la variation métrique intragroupe dans n'importe quel nombre d'échantillons et de créer un logiciel convivial à ces fins. Nous avons revisité le plan d'étude classique basé sur la comparaison entre les échantillons de Berg, Zalavar et Oslo de l'ensemble de données craniométriques de W.W. Howells. Quatorze dimensions du milieu du visage ont été choisies pour l'analyse. Le logiciel "WorldPCA" a été utilisé pour la plupart des analyses. Le logiciel met en œuvre un certain nombre de fonctions analytiques visant à explorer les résultats de l'analyse en composantes principales. Nos résultats ont confirmé que l'échantillon masculin de Berg présente un niveau de variation accru. Les analyses de grappes ont montré que la variation intragroupe dans les trois échantillons est principalement de nature continue. Nous pouvons dire que la tendance à se séparer en groupes distincts est plus prononcée dans les échantillons d'Oslo et de Berg que de Zalavar. Certains individus masculins de Zalavar présentent des caractéristiques craniofaciales distinctes, similaires à celles trouvées dans un échantillon du sud de la Sibérie. Les applications potentielles de ces techniques et logiciels ne se limitent pas aux mesures crâniennes mais peuvent être utilisées pour explorer tout type de données à variation continue. Aucune hypothèse sur la nature des données ne doit être faite, et n'importe quel nombre d'échantillons peut être comparé simultanément.

Mots clés – variation intragroupe, analyse en composantes principales, base de données Howells

Introduction

Intragroup variation in human cranial samples attracts much less attention from researchers and is much less well understood than intergroup variation (Stojanowski and Schillaci, 2006; Le Maître and Mitteroecker, 2020). This is not an exception in morphometric studies of different mammalian taxa but rather a common situation: the influence

of intragroup variation on intergroup biological distances is substantial yet poorly understood, particularly when phylogenetically close groups are compared and sample sizes are relatively small (Caumul and Polly, 2005; Cardini and Elton, 2007; Cardini et al., 2015). This holds true even more for modern *H. sapiens* displaying a number of specific factors that hugely affect the variability of our species' genome and thus its phenotype: rapid and continuous migrations and admixture, and demographic explosions (Jobling, 2012). This can be arguably complemented by relaxed selective pressure, at least for cranial form (Bunak, 1960; Roseman, 2016; Lacruz et al., 2019). As a result, the ratio of intra- to intergroup cranial variation has been repeatedly shown to be exceptionally high in *H. sapiens* (Relethford, 1994; 2002; Strauss and Hubbe, 2010), even compared to our closest primate relatives (Weaver, 2014).

However, during the last decades there have been relatively few studies directly addressing this issue (Le Maître and Mitteroecker, 2020; Howells, 1966; 1974; Cadien et al. 1974; Deryabin, 1998; Key and Jantz, 1990; Raemsch and Wilkinson, 1994; Petersen, 2000; Stefan, 2004; Shirobokov, 2018; 2020). Two of these studies focused on comparing the early Mediaeval Hungarian Zalavar sample with supposedly more homogenous samples from Mediaeval Oslo and an isolated Alpine village, Berg (Howells, 1974; Petersen, 2000). The first was exploratory in nature and based on the use of factor analysis, while the second employed a number of more sophisticated and strict statistical tests. Nevertheless, the results from both were similar in demonstrating the sample from Berg to be the most variable despite its cultural homogeneity and relative reproductive isolation. Virtually all the studies cited above employ one of the following approaches: a) comparison of morphological variation between a test sample and a reference ("etalon") sample; b) direct comparison of the degree of variation in samples of interest. In the first case, the very serious issue of choosing the reference sample arises (see Stefan, 2004) as neither cultural homogeneity nor familial relatedness are consistently associated with less cranial morphological variation (Howells, 1966; Raemsch and Wilkinson, 1994; Sardi et al., 2006; Evteev and Dvurechensky, 2017). In the second case, the comparison is basically restricted to just two samples (e.g. Petersen, 2000). Also, most of the approaches assume multivariate normality of the data or are sensitive to sample size (Petersen, 2000). In either case, the main outcome of these analyses is the finding that one sample is more variable than another, or that it is more (or less) variable compared to the reference sample. But such a conclusion, while important, does not answer a number of questions that are crucial for any assessment of intragroup variation:

- are individuals in a sample evenly distributed across the range (wide or narrow) of variation in the sample, or do they tend to form more or less distinct clusters?
- are morphological "outliers" present in the sample? These outliers can be representative of genetically and morphologically distinct populations or individuals displaying mild forms of craniofacial pathology;
- in which variables exactly is variation greater in the sample of interest? This question becomes important when considering the different "phylogenetic relevance" of different variables (Roseman, 2004; Betti et al., 2010; Evteev et al., 2020).

Principal component analysis (PCA) is a method for exploring intragroup variation which is capable of addressing the issues mentioned above. However, applying PCA to compare the intragroup structure of multiple samples means resolving a number of methodological problems (see Evteev and Dvurechensky, 2017 and https://sites.google.com/site/worldpcaeng/ for detail). First, the morphological meaning and variance of principal components will be strongly dependent on the sample's composition. As a possible solution, a large and diverse training sample can be employed to construct a PCA morphospace for a number of metric variables. Thus, a universal "background" for the comparison of intragroup variation in any number of samples can be created. This provides for objective assessment of the degree and pattern of intragroup variance in each of the samples of interest (Evteev and Dvurechensky, 2017:144). Another issue is the absence of suitable software for analytical explorations of the results of PCA analyses from the point of view of different aspects of intragroup variation.

The aims of this study were to develop a flexible and assumption-free approach for detailed explorations and comparisons of intragroup metric variation in any number of samples, and to create simple and user-friendly software for these purposes. As an illustration, we revisited the classic comparison between the samples from Berg, Zalavar and Oslo (Howells, 1974; Petersen, 2000).

Material and methods

Samples

The three samples – from Berg, Zalavar and Oslo – have been described in detail by Howells (1989:89-91), so we only provide a brief overview here, focusing on the aspects that might have influenced the degree of intragroup cranial variation.

Berg is a small mountain village near Greifenburg on the Drau River in Western Carinthia. At the time of acquisition (1911) the village had only about 100 inhabitants. The sample, supposedly representing all the crania of some five generations of this population, was first described by Shapiro (1929). According to Howells' visual observations, cradling practices were influential in shaping the cranial vault. He states that "Skulls... with no signs of secondary effects *(i.e. unintentional deformations – AE)*, were few" (Howells, 1989:91). Importantly, the skulls of this sample were quite difficult to sex, and the determination had to be based exclusively on visual cranial traits assessed by Howells only.

The sample from Mediaeval Oslo comes from a number of graveyards representing a more or less culturally homogenous population, but mainly (85%) from just one cemetery near the St. Nicolaus Church (Howells, 1989:89;

see also Schreiner, 1939). However, Howells does not provide any information on how long the cemetery was used. The selection of skulls based on the presence of the metopic suture was carried out where the cases of metopism were held to about 10% (Howells, 1989:89). Determination of sex was done by both Schreiner (1939) and his co-workers and later by Howells. Postcranial skeletons were available in some cases.

The Zalavar site is situated near Lake Balaton in western Hungary. It dates to the 9th-11th centuries AD and represents quite a complex population (Howells, 1989:90). The inhabitants of the Zalavar stronghold were extremely heterogeneous ethnically and included surviving Avars and some early Magyar, Germanic and Slavonic elements (the latter being probably the most numerous). In fact, the sample represents two different cemeteries at Zalavar that are likely to have been used by very different groups of people (Ascadi et al., 1962). A very thorough sex determination procedure, involving the postcranial skeleton, was carried out for this sample (Howells, 1989:90).

Thus, based on existing historical and archaeological knowledge, there is no doubt that Berg represents the most culturally homogenous community and Zalavar the most cosmopolitan one. But the cranial deformations observed and uncertainty in the sex determination might introduce additional sources of intragroup variation in the sample from Berg.

The training sample employed to construct a universal morphological "background" against which variation in the three samples was assessed included 24 (male) and 21 (female) world-wide populations from the Howells database (Howells, 1989; https://web.utk.edu/~auerbach/HOWL.htm) (table 1). Five of the samples from the original database were excluded from the analysis: Bushman, two samples of Australian aborigines ("Australia" and "Tolai"), Tasmania and Andaman. This was done because the very specific craniofacial morphology of these groups – either of genetically determined very small body size (Bushman and Andaman) or isolated (Australia and Tasmania) could have disproportionately affected the results of the analyses. Two Maori samples (southern and northern) were combined into one because of the small sample size.

The male (1132 individuals altogether) and female (926 individuals) skulls were analyzed separately (table 1).

Variables

Of 82 linear measurements and angles that can be found in the Howells database, only 14 mid-facial dimensions were chosen for the analysis: NLH, NLB, MAB, OBH, OBB, DKB, NDS, WNB, SIS, ZMB, SSS, XML, MLS, WMH. The reason for excluding cranial vault measurements is twofold. First, in general, the influence of a number of social and environmental factors, as well as secular trends, produces rapid changes in the shape of the cranial vault, which makes the comparison and interpretation of diachronic cranial samples difficult in many cases (Debets, 1948; Alekseeva, 1973; Beals et al., 1984; Jantz et al., 2000; Godina et al. 2005). The mid-facial skeleton is expected to be more stable in this sense (Bunak, 1960; Lacruz et al., 2019; Evteev et al., 2020). Second, as noted above, unintentional cranial deformations were present in virtually all the skulls from Berg, which could seriously affect the pattern of intragroup variation in this sample. We decided to focus our analyses exclusively on mid-facial dimensions as these have been consistently shown to be good indicators of population history: the association between intergroup craniometric and genetic distances based on mid-facial measurements is as high as for the cranium in general (see for review Evteev et al., 2020; von Cramon-Taubadel, 2014).

Population[1]	Number of male skulls	Number of female skulls	Total
Norse	55	55	110
Zalavar	53	45	98
Berg	56	53	109
Egypt	58	53	111
Teita	33	50	83
Dogon	46	52	98
Zulu	55	46	101
Mokapu	51	49	100
Buryat	54	54	108
Inuit[2]	53	55	108
Peru	55	55	110
Arikara	42	27	69
Ainu	47	38	85
North Japan	55	32	87
South Japan	50	41	91
Hainan	45	38	83
Anyang	42	0	42
Atayal	29	18	47
Philippines	50	0	50
Guam	30	27	57
Easter Island (I)	48	36	84
Moriori	56	51	107
Maori (South and North)	18	0	18
Santa Cruz	51	51	102
Total	1132	926	2058

1 – names of the populations according to the Howells dataset;
2 – this population is entitled "Eskimo" in the Howells dataset.

Table 1. Cranial sample used in this study / *Échantillon crânien utilisés dans cette étude*

Statistical analyses

The "WorldPCA" software written by one of the authors (ES) using Python 3.6 with the SciPy, Matplotlib, scikit-learn and Pandas libraries (Hunter, 2007; McKinney, 2010; Millman and Aivazis, 2011; Pedregosa et al., 2011) was employed for most of the analyses carried out in this study (https://github.com/Nikolay-Staroverov/WorldPCA, where manuals and the individual craniometric data employed in this study can be found as well; please visit https://drive.google.com/drive/folders/1C6g4ckwCIOlnMGBX4aqMMLEPX-8KHcz5V?usp=sharing to download an executable version of the software). The software performs principal component analysis (PCA) based on correlation matrices where the individuals analyzed can be divided into training and test sets. The PC equations are calculated using only the training set, and coefficients from those equations are used to calculate PC values for the individuals included in the test set. In this study, however, the sample was not divided into two datasets but was analyzed as a whole since the samples from Oslo, Berg and Zalavar are the only West Eurasian samples in the Howells database and their exclusion would have biased the training set.

The analysis results can be visualized in WorldPCA in the form of scatterplots and diagrams. But, more importantly, the software implements a number of analytical tools aimed at exploring the results of the PCA. Virtually all of these functions are based on a very simple statistic: mean pairwise Euclidean distance (MPED) which can be calculated between individuals in the same sample or different samples using various combinations of PCs. The functions implemented in WorldPCA are as follows:

1) Search for individuals that can be called "morphological outliers", i.e. those lying well outside the range of the sample's morphological variation. This is carried out by comparing the MPED of the individual from other individuals in the sample with the mean MPED of this sample. The output is presented in the form of bar charts.
2) Assessment of the total range of variation in a sample (mean MPED) in the morphospace of chosen PCs. This range can be compared with the ranges of other samples, and the difference can be tested for significance using the Mann-Whitney U-test.
3) Search for clusters of individuals in samples using different methods of clustering: k-means, Birch, agglomerative clustering. Individuals can also be assigned to clusters manually. The software calculates the MPED of individuals in a particular cluster from other individuals in the sample (i.e. not from this cluster). When compared with the MPED of the whole sample, this can demonstrate how morphologically distinct the cluster is.
4) Intergroup MPED can be calculated as well.

In order to simplify the presentation of the results, only the first four PCs are discussed. Figures 1 and 3 were prepared in Excel based on the intragroup MPEDs calculated in WorldPCA (using modules "Distance" and "Clustering", respectively). Figure 6 was drawn in PAST (Hammer et al., 2001) based on PC scores calculated in WorldPCA.

Results

Variation in the total sample

The PC structures of the analyses of the male and female world-wide samples are fairly similar (table 2). The PC loadings of the two analyses are moderately to highly correlated (Spearman R): 0.93 ($p<0.0001$), 0.35 ($p=0.23$), 0.78 ($p=0.001$), and 0.97 ($p<0.0001$)* for PC1, 2, 3 and 4, respectively. The only relatively low and non-significant coefficient was obtained for the second PC, although its morphological meaning is nevertheless almost identical in the male and female samples (see table 2). PC1 exhibits eigenvalues of 3.5/3.43 (males/females) and describes 25/24.5% of the total variance. High PC1 scores are associated with an overall increase in the nasal, orbital and malar sizes with no notable change in the nasal bone and nasal bridge dimensions (DKB, NDS, WNB, SIS, and also ZMB). PC2 has eigenvalues of 2.2/2.2 and describes 15.6/15.6% of the total variance. High PC2 scores will correspond to individuals with larger dimensions of the nasal bones and bridge (see above). The eigenvalues of PC3 are 1.8/2.0; this describes 13.1/14.0% of the total variance. High PC3 scores will be observed in individuals displaying a high and narrow nasal aperture, tall orbits and a narrow nasal bridge (DKB). Finally, PC4 exhibits eigenvalues of 1.2/1.3 and describes 8.6/8.9% of the total variance. It is mainly associated with a decrease in subnasal protrusion (SSS) accompanied by an increase in malar dimensions (XML and MLS) and cheek height (WMH). The cumulative percentage of the total variance explained by the first four PCs is 62.2% for the male sample and 63.1% for the female sample. None of the lower PCs exhibits an eigenvalue higher than 1 and were therefore not considered further, in accordance with Kaiser's rule (Jackson, 1993).

Intragroup variation in the Zalavar, Berg and Oslo samples at the worldwide scale

A comparison of intragroup MPEDs calculated using all four PCs showed (figure 1) that the samples from Oslo and Zalavar, both male and female, display moderate levels of intragroup variation of the mid-facial shape relative to other samples from the Howells database. But the situation is different for Berg: while females from this village are moderately variable, males demonstrate the second-highest level of intragroup variation among all the samples.

Clustering

Each of the three samples was divided into three clusters using the k-means method (figures 2 and 3). Note that the clustering procedure employed all four PCs while only PC1 and PC2 are presented in the plots. Clearly, none of the

Variable	PC1		PC2		PC3		PC4	
	M	F	M	F	M	F	M	F[1]
Principal components loadings								
NLH	0.68	0.59	0.20	-0.04	0.35	0.54	-0.06	-0.05
NLB	0.35	0.29	0.10	0.59	-0.60	-0.33	-0.36	-0.35
MAB	0.68	0.64	0.08	0.24	-0.14	0.12	-0.17	-0.15
OBH	0.48	0.44	-0.01	-0.32	0.53	0.44	-0.17	-0.24
OBB	0.58	0.54	0.14	-0.06	0.33	0.33	-0.08	0.03
DKB	0.05	-0.10	0.32	0.76	-0.74	-0.27	-0.14	-0.22
NDS	-0.16	-0.40	0.77	0.44	0.19	0.56	0.11	0.23
WNB	-0.17	-0.36	0.75	0.73	-0.33	0.16	0.15	0.13
SIS	-0.19	-0.40	0.83	0.34	0.22	0.68	0.21	0.31
ZMB	0.75	0.63	0.03	0.40	-0.24	-0.01	-0.18	-0.31
SSS	0.13	-0.07	0.33	0.08	0.20	0.48	-0.62	-0.48
XML	0.69	0.67	0.07	0.17	-0.02	-0.03	0.44	0.48
MLS	0.53	0.56	-0.21	0.21	-0.29	-0.34	0.49	0.50
WMH	0.67	0.69	0.12	0.07	0.02	0.14	0.23	0.14
Eigenvalues and related statistics								
Eigenvalue	3.5	3.4	2.2	2.2	1.8	2.0	1.2	1.2
% of total variance	25.0	24.5	15.6	15.6	13.1	14.0	8.6	8.9
% of total variance, cumulative	25.0	24.5	40.6	40.1	53.7	54.2	62.2	63.1

1 - the loadings for the same variables had different signs in the male and female analyses, and for the latter were multiplied by -1

Table 2. Results of principal component analyses of the world-wide male and female samples / *Résultats de l'analyse en composantes principales de l'échantillon international de sujets féminins et masculins*

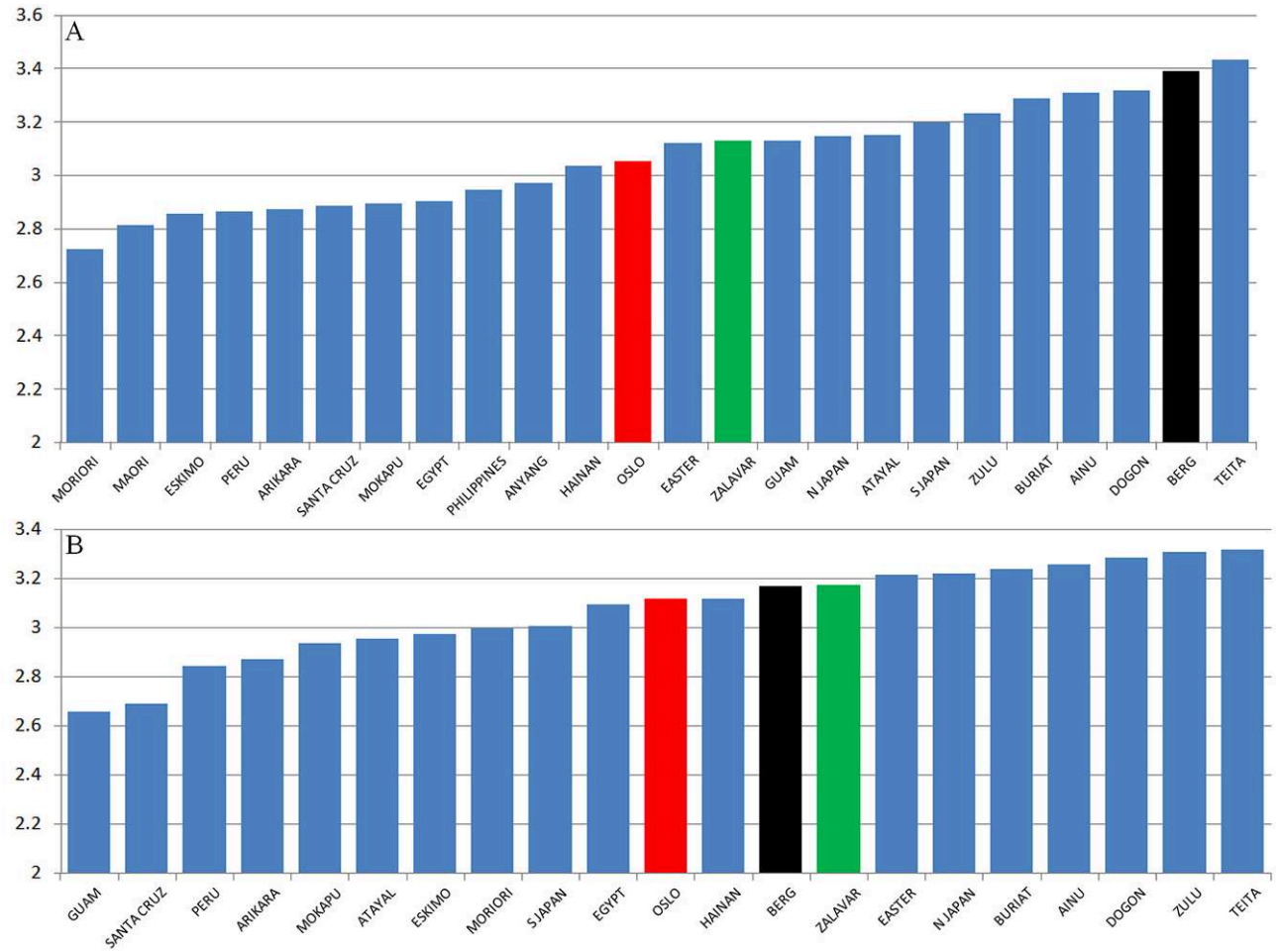

Figure 1. Intragroup MPED (y-axis) for PC1-4. (a) male samples. (b) female samples / *Distance euclidienne intragroupe par paire (MPED) pour les composantes principales 1-4. (a) échantillon masculin. (b) échantillon féminin*

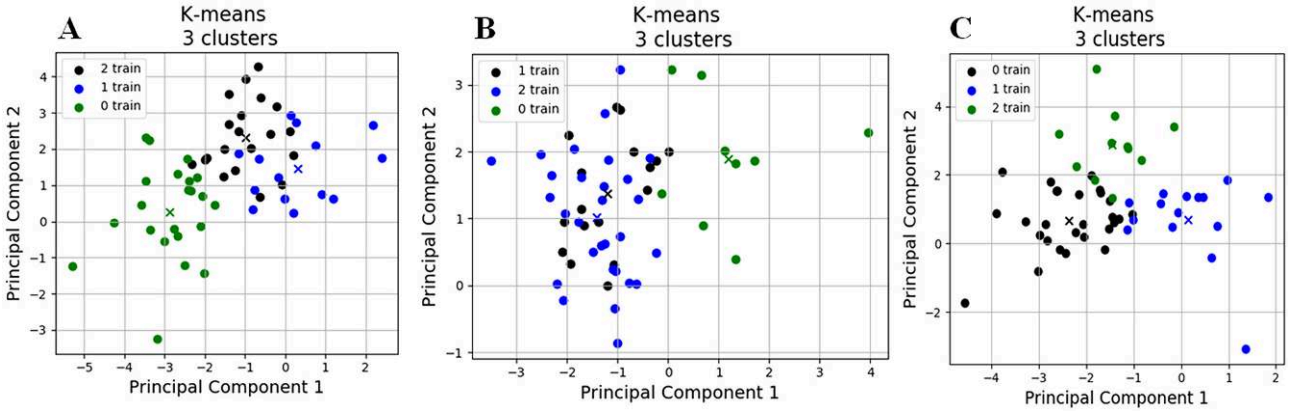

Figure 2. Clusters in the male samples from Berg (a), Oslo (b) and Zalavar (c). Crosses represent the means of the respective clusters. Note, that the clustering procedure employed all four PCs while only PC1 and PC2 are shown in the plots / *Groupes d'échantillons masculins de Berg (a), d'Oslo (b) et de Zalavar (c). Les croix représentent les moyennes des groupes respectifs. Il faut noter que la procédure de regroupement utilise l'ensemble des quatre composantes principales (CP) alors que seules la CP1 et la CP2 sont représentées sur le graphique*

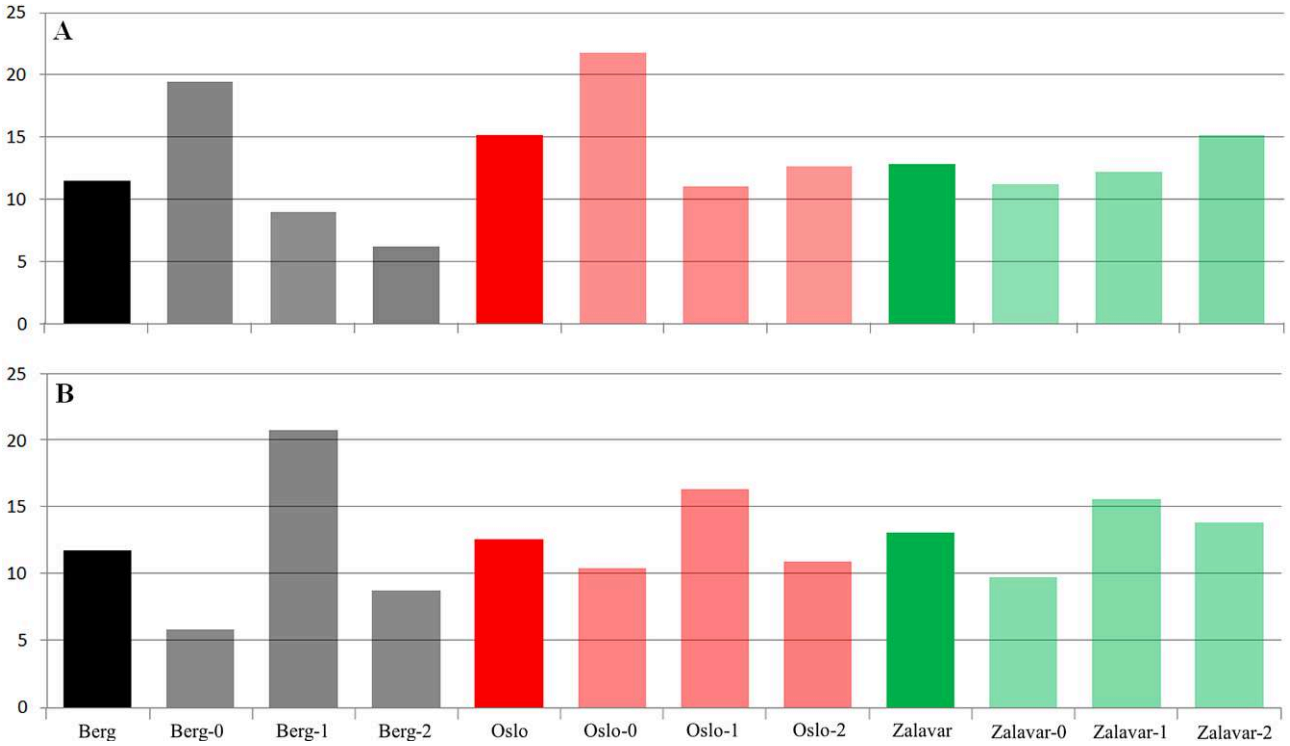

Figure 3. Mean pair-wise Euclidean distance between individuals from a cluster and other individuals in a sample (see text for details). (a) male samples. (b) female samples / *Distance euclidienne par paire entre les individus d'un groupe et ceux d'un échantillon (voir le texte pour plus d'informations). (a) échantillon masculin. (b) échantillon féminin*

samples display segregation of well defined clusters but rather a continuous variation. Nevertheless, MPEDs for some clusters were substantially higher than the MPEDs of the respective samples (figure 3). In figure 3, the light-coloured bars stand for the difference (in %) between the MPEDs of a sample and of a cluster. For instance, in the male sample from Berg, the MPED between all individuals averages 3.39 while the MPED between the individuals belonging to cluster "0" (figure 2a) and other individuals in the sample is 4.05, i.e. 19.4% higher. The dark-coloured bars represent the mean MPED for the three clusters in each of the samples. Based on the results of this analysis, the sample from Berg can be again considered the most variable as clusters with relatively high MPEDs were detected in this sample. These are clusters "0" in males (19.4% higher than average MPED) and "1" in females (20.8%). The male sample from Oslo also exhibits clusters with high MPEDs, e.g. cluster "0" (21.8%). The tendency to form distinct morphological clusters in the sample from Zalavar is at least not stronger in comparison to the two other samples.

Assessment of the proportion of individuals substantially deviating from the sample mean: intragroup aspect

The number of individuals lying outside the main range of morphological variation of a sample can vary substantially between different groups (figures 4-5). A possible criterion for defining such individuals can be an individual MPED one standard deviation (SD) higher than the MPED of a sample (upper red dashed lines in figures 4 and 5). Based on this criterion, the following proportions are observed in the samples: Berg -17.9%/18.9% (males/females); Oslo -10.9%/10.6%; Zalavar -17.0%/15.6%. Thus, the samples from Berg and Zalavar are almost equally "heterogeneous" while the sample from Oslo is substantially more "homogenous". An interesting finding is the presence of strongly deviating individuals (outliers) in some samples: "84" in the male sample from Zalavar, "41" in the female sample from Oslo or "112" in the female sample from Berg.

Assessment of the proportion of individuals substantially deviating from the sample mean: intergroup aspect

One of the reasons why the cranial sample from Zalavar is expected to display increased morphological variation is the possible presence of representatives of geographically very remote human groups of Central Asian, East Asian and Uralian origin: Avars and early Magyar. Only one sample in the Howells dataset can be used to test this hypothesis, namely the Buryats ("Buriat" according to Howells). The Buryats are a Mongolian-speaking ethnic group living mainly to the south of Lake Baikal and closely connected both historically and anthropologically with the nomadic populations of Eastern Central Asia [45]. A comparison of the male samples from Oslo, Berg and Zalavar with the Buryats shows (figure 6a) a very small overlap between the former and the latter, mainly for the PC1 values between 0 and 2.5 and PC2 values in the range of -0.5 to 2.5. A very similar

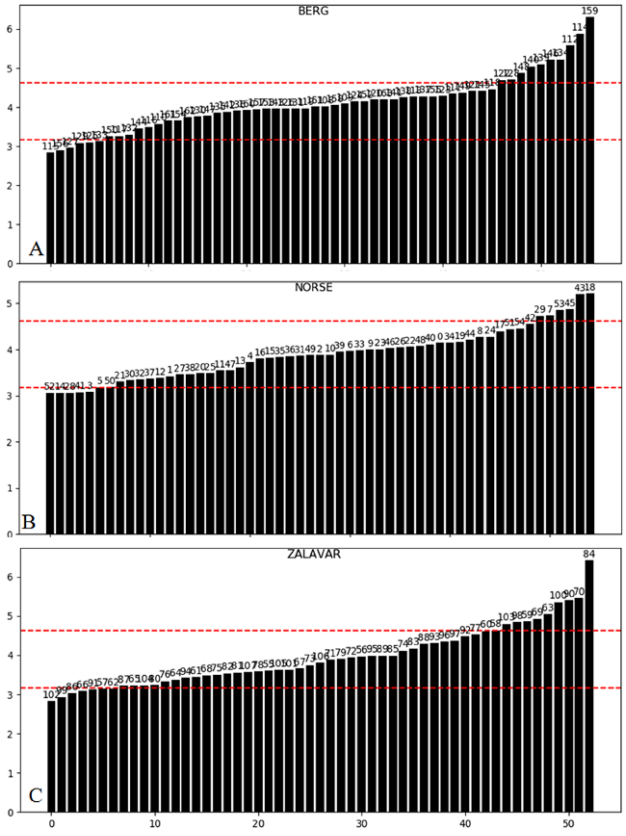

Figure 4. Distribution of individual MPEDs in the samples from Berg (a), Oslo (b) and Zalavar (c). Males. Lower dashed line – mean MPED, upper dashed line – mean MPED + 1 standard deviation. Individuals are arranged in ascending order along the x-axis according to their respective MPEDs / *Distribution des distances euclidiennes par paire (MPED) dans les échantillons de Berg (a), Oslo (b) et Zalavar (c) pour les individus masculins. Ligne pointillée du bas – moyenne des MPED, ligne pointillée du haut – moyenne des MPED + 1 écart-type*

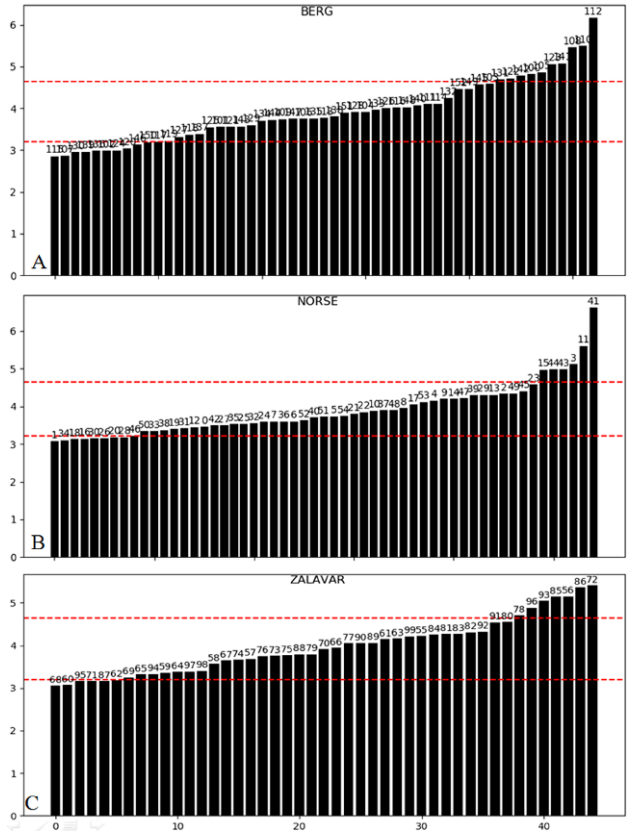

Figure 5. Distribution of individual MPEDs in the samples from Berg (a), Oslo (b) and Zalavar (c). Females. Lower dashed line – mean MPED, upper dashed line – mean MPED + 1 standard deviation. Individuals are arranged in ascending order along the x-axis according to their respective MPEDs / *Distribution des distances euclidiennes par paire (MPED) dans les échantillons de Berg (a), Oslo (b) et Zalavar (c) pour les individus féminins. Ligne pointillée du bas – moyenne des MPED, ligne pointillée du haut – moyenne des MPED + 1 écart-type*

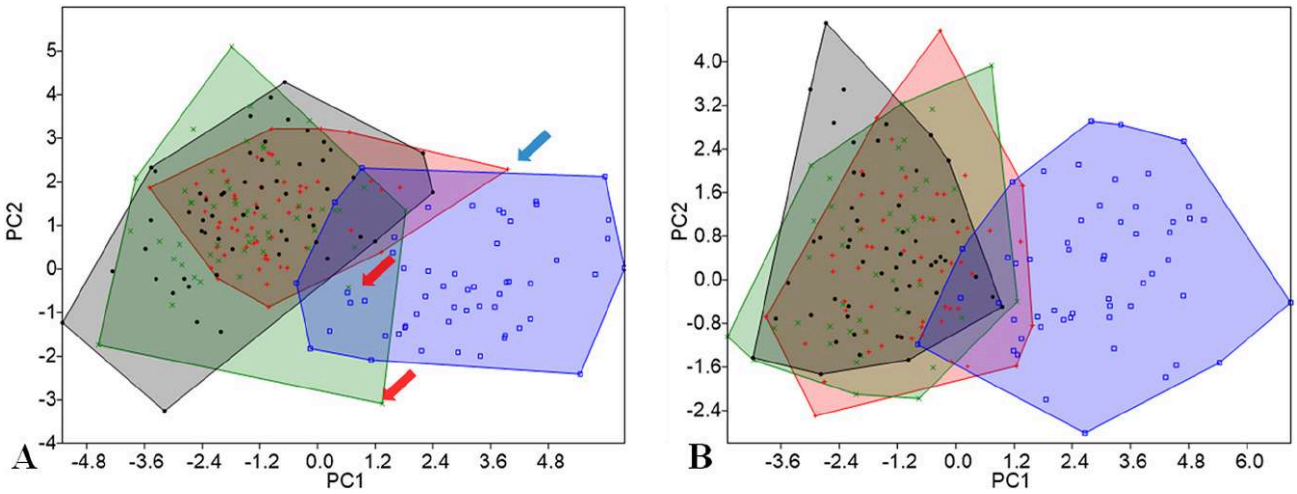

Figure 6. Scores for PC1 and PC2 for the samples from Berg (black), Oslo (red), Zalavar (green) and the Buryat sample (blue). (a) males. (b) females / *Résultats des composantes principales 1 et 2 pour les échantillons de Berg (noir), Oslo (rouge), Zalavar (vert) and de Buryat (Bleu). (a) individus masculins. (b) individus féminins*

picture is observed for females (figure 6b) whereas the area of overlap spans approximately -1.0 to 1.5 of PC1 and -1.6 to 2.0 of PC2. However, in the male sample from Zalavar there are two individuals (red arrows) falling outside the common area of overlap between the European and Buryat samples and displaying PC1 and PC2 values typical for the Buryats. One individual from Oslo displays an exceptionally high PC1 value (blue arrow). There are no such individuals in the female parts of the same samples.

Discussion

The aims of this study were to develop a PCA-based approach for studying and comparing intragroup variation in different samples and to create simple and user-friendly software for this purpose. In order to test these, we compared three European samples from the Howells database: Berg, Zalavar and Oslo (Howells, 1989). These three samples have been previously extensively studied from the point of view of intragroup variation using different statistical techniques (Howells, 1974; Petersen, 2000). Such a sample choice makes it is possible to directly compare the results of the previous and present studies and to evaluate the relative efficacy of different approaches to understanding the degree and pattern of intragroup variation in these samples.

Note that only mid-facial craniometric variables were included in the analyses and discussed further (see Introduction for the rationale behind this choice).

The degree of intragroup variation in the samples of interest was first considered in a world-wide context by comparing them with other groups included in the Howells data set. In previous studies, this was only compared between the three samples or with a supposedly homogeneous "etalon" sample (Howells, 1974; Petersen, 2000). Our results have confirmed that the male sample from Berg displays a higher degree of variation – not only with respect to the samples from Oslo and Zalavar, but also compared to all but one population from the data set (figure 1a). The Norwegian and Hungarian samples, in turn, display intermediate rather than low levels of variation (see figure 1). Two important new conclusions can be drawn from these particular results. First, the high degree of variation in males from Berg is not due to the unintentional cranial deformations observed in the bulk of the individuals of this sample (Howells, 1989:91) since it has been shown that the morphological changes occurring as a result of cranial vault deformations only affect peripheral structures of the facial skeleton (Anton, 1989). Second, the females from Berg display a smaller degree of variation than males of the same sample, unlike the females from Oslo and Zalavar. Interestingly, Berg is one of a few samples (also Guam and Atayal) where males and females are very different in terms of the degree of intragroup variation whereas in most other cases no particular sexual dimorphism is observed. This might reflect aspects of the population and demographic history of these groups that would deserve more detailed consideration.

The cluster analyses have shown that intragroup variation in the three samples is mainly of a continuous nature. But some clusters in some samples tend to be more distinct than others. Arguably, the tendency to separate into such clusters is more pronounced in the samples from Oslo and Berg than from Zalavar, the most distinct clusters in the first and second at least being more prominent. It is difficult to develop a strict significance test for the clustering procedure as the MPED of virtually any objectively defined cluster occupying a particular area of the morphospace will be clearly different from the MPED for a whole sample. Rather, the degree of "distinctness" of a cluster should be determined in a comparative context, relative to other clusters of the same or other samples. Though not fully statistically strict, clustering can provide interesting information about

patterns of intragroup variation in a sample. For instance, the most distinct clusters in the male samples from Oslo and Zalavar display increased values of PC1 scores, i.e. a tendency to an increase in general facial size with no notable change in the nasal bone and nasal bridge dimensions, whereas the most distinct cluster in the male sample from Berg is characterized by a simultaneous decrease in PC1 and PC2 values. Individuals in this cluster thus display a smaller face but also narrower and less protruding nasal bones.

The presence of particularly morphologically peculiar or atypical individuals was assessed using bar charts depicting the MPED of individuals from all other individuals in a sample. The steepness of such bar charts (e.g. evaluated by the broken stick criterion; Jackson, 1993) can also be used as an indicator of the degree of intragroup variation. The application of this criterion again points towards an increased variation in the male sample from Berg. Another aspect is the presence of individuals exhibiting very high MPEDs: "84" for the male sample from Zalavar, "41" for the female sample from Oslo, "112" for the female sample from Berg. There could be various reasons for such morphological distinctiveness, including subtle craniofacial disorders and observer errors. Special attention should be paid to such individuals and, in some cases, it may be reasonable to exclude them from the sample. The most interesting possible reason could be that such "atypical" individuals are representative of morphologically distinct and geographically remote populations. But in order to confirm this, a specific similarity to a particular population would need to be demonstrated.

In this study, we were able to test whether some individuals from Zalavar are atypical for their sample but display a similarity to the Buryats, a Mongolian-speaking group of Central Asian origin. Such similarity can be expected based on the knowledge of the origin of part of the early Mediaeval population of Zalavar (Howells, 1974; Ascadi et al. 1962). Our analyses show that in general, the three European samples only marginally overlap with the Buryat sample. But in the male sample from Zalavar, two individuals show combinations of PC1 and PC2 values not observed in the samples from Oslo or Berg, putting them into the range of morphological variation of the Buryats. Importantly, in the female sample of Zalavar such individuals are not found. This is not surprising since the long-range westward migrations of Asian nomadic tribes in Mediaeval times are known to be strongly male-biased (e.g. Ascadi et al., 1962). This observation at once corroborates and contradicts the conclusions of Howells (Howells, 1974): on the one hand, Howells interprets the combination of subnasal flatness and lack of prognathism in the sample from Zalavar as a result of "some slight Avar presence" (Howells, 1974:93), which is in line with the findings of the present study; on the other hand, according to Howells, it is impossible to single out any "types" or individual variants reflecting the nature of ancestral groups of this population. This view cannot be completely accepted based on the results of our analyses: while the male sample from Zalavar in general clearly exhibits West Eurasian cranial morphology and, moreover, is quite homogeneous, some individual skulls display distinct features similar to the South Siberian reference group.

We consecutively considered different possible manifestations of heterogeneity of a sample: the presence of clusters, morphologically atypical individuals or an increased range of continuous variation. A sample can display only one of these aspects or several of them simultaneously. The software presented in this study provides the means to visualize these different types of intragroup variation and to separate them at least qualitatively.

Our concluding remarks relate to potential applications of the techniques and software employed in this study. These are not restricted to cranial measurements but can be used to explore any type of continuously varying data. In contrast to some existing methods of analyzing intergroup variation (Key and Jantz, 1990; Stefan, 2004; Shirobokov, 2018; 2020; but see Petersen, 2000), no particular assumptions about the nature of the data should be made, and any number of samples can be compared simultaneously instead of just one (i.e. reference vs. etalon) or two. As the approach is based on principal component analysis, it is well suited for assessing not only differences but also similarities between samples. Of course, analyzing many samples together may not produce a full description of each nuance of intragroup variation in particular groups, but this is compensated by an important opportunity for making the patterns of this variation fully comparable and interpretable.

Acknowledgements

This study was supported by the Ministry of Science and Higher Education of the Russian Federation, system number 075-10-2020-116 (grant number 13.1902.21.0023).

References

Alekseeva TI (1973) Etnogenez vostochnykh slavyan po dannym antropologii, MSU Publishing, Moscow, 330 p [in Russian]

Alexeev VP, Gohman II (1984) Antropologiya aziatskoj chasti SSSR, Nauka, Moscow, 207 p [in Russian]

Anton SC (1989) Intentional cranial vault deformation and induced changes of the cranial base and face. American Journal of Physical Anthropology 79:253-267

Ascadi G, Harsanyi L, Nemeskery G (1962) The population of Zalavar in the Middle Ages. Acta Archaeologica Academiae Scientiarum Hungaricae 14:113-141

Beals KL, Smith CL, Dodd SM (1984) Brain size, cranial morphology, climate and time machines. Current Anthropology 25: 301-330

Betti L, Balloux F, Hanihara T, Manica A (2010) The relative role of drift and selection in shaping the human skull. American Journal of Physical Anthropology 141:76-82

Bunak VV (1960) Litsevoi skelet i faktori, opredelyayushchiye variatsii ego stroeniya. In: Debets GF, Levin MG (eds) Proceedings of the USSR Academy of Sciences. USSR Academy of Science, new series, vol L. Moscow, pp 84-152 [in Russian]

Cadien JD, Harris EF, Jones WP, Mandarino LJ (1974) Biological lineages, skeletal populations, and microevolution. Yearbook of Physical Anthropology 18:194-201

Cardini A, Elton S (2007) Sample size and sampling error in geometric morphometric studies of size and shape. Zoomorphology 126:121-134

Cardini A, Seetah K, Barker G (2015) How many specimens do I need? Sampling error in geometric morphometrics: testing the sensitivity of means and variances in simple randomized selection experiments. Zoomorphology 134:149-163

Caumul R, Polly PD (2005) Phylogenetic and environmental components of morphological variation: skull, mandible and molar shape in marmots (Marmota, Rodentia). Evolution 59: 2460-2472

von Cramon-Taubadel N (2014) Evolutionary insights into global patterns of human cranial diversity: population history, climatic and dietary effects. Journal of Anthropological Sciences 92: 43-77

Debets GF (1948) Paleoantropologiya SSSR, AS USSR Publishing, Moscow/Leningrad, 389 p [in Russian]

Deryabin VE (1998) Razdelenie neodnorodnikh vyborok v kraniologii: mnogomernyi podhod. Voprosy antropologii 89:36-50 [in Russian]

Evteev AA, Dvurechensky OV (2017) Intra-Group Variation of the Facial Skeleton in 16th-19th Century Rural Russian Populations in the Worldwide Context: A Principal Component Analysis. Archaeology, Ethnology and Anthropology of Eurasia 45: 143-151

Evteev A, Santos P, Ghirotto S, Reyes-Centeno H (2020) Associations between genetic differentiation and craniometric interpopulation distances across North Eurasia. American Journal of Physical Anthropology 171(S69):82

Godina EZ, Khomyakova IA, Purundzhan AL, Zadorozhnaya LV, Gilyarova OA, Stepanova AV (2005) Moskovskie deti: osnovnye tendetsii rosta I razvitiya na rubezhe stoletij. Chast 2. Voprosy antropologii 92:56-75

Hammer Ø, Harper DAT, Ryan PD (2001) PAST: Paleontological statistics software package for education and data analysis, [www.nhm.uio.no/english/research/infrastructure/past]

Howells WW (1966) Variability in family lines vs. population variability. Annals of the New York Academy of Sciences 134: 624-631

Howells WW (1974) The population of Zalavar: a problem in cranial variation. Anthropologiai Közlemények 18:91-96

Howells WW (1989) Skull shapes and the map. Peabody Museum Press, Cambridge, USA, 187 p

Hunter J (2007) Matplotlib: A 2D Graphics Environment. Computing in Science and Engineering 9:90-95

Jackson DA (1993) Stopping Rules in Principal Components Analysis: A Comparison of Heuristical and Statistical Approaches. Ecology 74:2204-2214

Jantz RL, Meadows Jantz L (2000) Secular change in craniofacial morphology. American Journal of Human Biology 12:327-338

Jobling MA (2012) The impact of recent events on human genetic diversity. Philosophical Transactions of the Royal Society B: Biological Sciences 367:793-799

Key PJ, Jantz RL (1990) Statistical measures of intrasample variability. Human Evolution 5:457-469

Lacruz RS, Stringer CB, Kimbel WH, Wood B, Harvati K, O'Higgins P, Bromage TG, Arsuaga JL (2019) The evolutionary history of the human face. Nature Ecology & Evolution 3:726-736

Le Maître A, Mitteroecker P (2020) Comparison of covariation patterns in craniofacial morphology across human populations. Bulletins et Mémoires de la Société d'Anthropologie de Paris 32(S9-33):21

McKinney W (2010) Data Structures for Statistical Computing in Python. In: van der Walt S, Millman J (eds) Proceedings of the 9th Python in Science Conference. Austin, Texas, pp 56-61 [https://conference.scipy.org/proceedings/scipy2010]

Millman KJ, Aivazis M (2011) Python for Scientists and Engineers. Computing in Science & Engineering 13:9-12

Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay E (2011) Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research 12:2825-2830

Petersen HC (2000) On statistical methods for comparison of intrasample morphometric variability: Zalavar revisited. American Journal of Physical Anthropology 113:79-84

Raemsch CA, Wilkinson RG (1994) Intrasample craniometric variability in related and unrelated skeletal samples. American Journal of Physical Anthropology 18:164

Relethford JH (1994) Craniometric variation among modern human populations. American Journal of Physical Anthropology 95:53-62

Relethford JH (2002) Apportionment of global human genetic diversity based on craniometrics and skin color. American Journal of Physical Anthropology 118:393-398

Roseman CC (2004) Detecting interregionally diversifying natural selection on modern human cranial form by using matched molecular and morphometric data. Proceedings of the National Academy of Sciences of the United States of America 101: 12824-12829

Roseman CC (2016) Random genetic drift, natural selection, and noise in human cranial evolution. American Journal of Physical Anthropology 160:582-592

Sardi ML, Novellino PS, Pucciarelli HM (2006) Craniofacial Morphology in the Argentine Center-West: Consequences of the Transition to Food Production. American Journal of Physical Anthropology 130:333-343

Schreiner KE (1939) *Crania norvegica*, H. Aschehoug & Co, Oslo, 388 p

Shapiro HL (1929) Contributions to the craniology of Central Europe. I. Crania from Greifenberg in Carinthia, Anthropological papers of the American Museum of Natural History 31 (1), New York, 124 p

Shirobokov IG (2018) Veroyatnostnyi podkhod k otsenke gipotez I problema mnozhestvennykh sravneniy v kraniologii. Kunskamera 1:87-96 [in Russian]

Shirobokov IG (2020) O primenenii metoda glavnykh komponent v kraniologii. Camera praehistorica 1:121-138 [in Russian]

Stefan VH (2004) Assessing intrasample variation: Analysis of Rapa Nui (Easter Island) museum cranial collections example. American Journal of Physical Anthropology 124:45-58

Stojanowski CM, Schillaci MA (2006) Phenotypic approaches for understanding patterns of intracemetery biological variation. American Journal of Physical Anthropology 131:49-88

Strauss A, Hubbe M (2010) Craniometric Similarities Within and Between Human Populations in Comparison With Neutral Genetic Data. Human Biology 82:315-330

Weaver TD (2014) Brief communication: quantitative and molecular-genetic differentiation in humans and chimpanzees: implications for the evolutionary processes underlying cranial diversification. American Journal of Physical Anthropology 154: 615-620